Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/IL05/000329

International filing date: 23 March 2005 (23.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/555,333
Filing date: 23 March 2004 (23.03.2004)

Date of receipt at the International Bureau: 27 May 2005 (27.05.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

PCT/IL2005/000329
18 MAY 2005

PA 1302241

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

April 05, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: *60/555,333*
FILING DATE: *March 23, 2004*

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

W. Montgomery

W. MONTGOMERY
Certifying Officer

13281 U.S. PTO

3

PTO/SB/16 (08-03)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

22264 U.S. PTO 60/555333

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No.

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
RAM	ZAMIR	TEL-AVIV, ISRAEL
ISHAI	ILANI	DOLEV, ISRAEL

Additional inventors are being named on the ______ separately numbered sheets *attached hereto*

TITLE OF THE INVENTION (500 characters max)

BASIS SELECTION IN A FINITE SET

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☒ Customer Number: 24505

OR

☐ Firm or Individual Name	ALPHAPATENT ASSOCIATES LTD.				
Address	P.O.B. 2345				
Address					
City	BEIT SHEMESH	State		Zip	99544
Country	ISRAEL	Telephone	516-620-4572 (US)	Fax	800-243-2384 (US)

ENCLOSED APPLICATION PARTS (check all that apply)

☒ Specification *Number of* Pages 21
☐ CD(s), Number ______
☐ Drawing(s) *Number of* Sheets ______
☐ Other (specify) ______
☐ Application Date Sheet. See 37 CFR 1.76

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ Applicant claims small entity status. See 37 CFR 1.27.
☐ A check or money order is enclosed to cover the filing fees.
☒ The Director is herby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 501380
☐ Payment by credit card. Form PTO-2038 is attached.

FILING FEE Amount ($): $80.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.
☐ Yes, the name of the U.S. Government agency and the Government contract number are: ______

Respectfully *submitted,*

[Page 1 of 2]

SIGNATURE ______

Date March 23, 2004

TYPED or PRINTED NAME DANIEL J. SWIRSKY

REGISTRATION NO. 45,148
(if appropriate)
Docket Number: 1310-USP1

TELEPHONE 516-620-4572 (US)

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

PTO/SB/17 (10-03)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

FEE TRANSMITTAL for FY 2004

Effective 10/01/2003. Patent fees are *subject to* annual revision.

☒ Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT	($) 80.00

Complete if Known	
Application Number	
Filing Date	MARCH 23, 2004
First Named Inventor	Ram ZAMIR
Examiner Name	
Art Unit	
Attorney Docket No.	1310-USP1

METHOD OF PAYMENT (*check all* that apply)

☐ Check ☐ Credit card ☐ Money Order ☐ Other ☐ None

☒ Deposit Account:

Deposit Account Number	501380
Deposit Account Name	ALPHAPATENT ASSOCIATES LTD.

The Director is authorized to: (*check all* that apply)

☒ Charge fee(s) indicated below ☒ Credit any overpayments

☒ Charge any additional fee(s) or any underpayment of fee(s)

☐ Charge fee(s) indicated below, except for the filing fee

to the above-identified deposit account.

FEE CALCULATION

1. BASIC FILING FEE

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description	Fee Paid
1001	770	2001	385	Utility filing fee	80.00
1002	340	2002	170	Design filing fee	
1003	530	2003	265	Plant filing fee	
1004	770	2004	385	Reissue filing fee	
1005	160	2005	80	Provisional filing fee	

SUBTOTAL(1) ($) 80.00

2. EXTRA CLAIM FEES FOR UTILITY AND REISSUE

			Extra Claims		Fee from below	Fee Paid
Total Claims		-20- =		X		
Independent Claims		- 3- =		X		
Multiple Dependent						

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description
1202	18	2202	9	Claims in excess of 20
1201	86	2201	43	Independent claims in excess of 3
1203	290	2203	145	Multiple dependent claim, if not paid
1204	86	2204	43	** Reissue independent claims over original patent
1205	18	2205	9	** Reissue claims in excess of 20 and over original patent

SUBTOTAL (2) ($)

**or number previously paid, if greater; For Reissues, see above

FEE CALCULATION (continued)

3. ADDITIONAL FEES

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description	Fee Paid
1051	130	2051	65	Surcharge -late filing fee or oath	
1052	50	2052	25	Surcharge - late provisional filing fee or cover sheet	
1053	130	1053	130	Non-English specification	
1812	2,520	1812	2,520	For filing a request for ex parte reexamination	
1804	920*	1804	920*	Requesting publication of SIR prior to Examiner action	
1805	1,840*	1805	1,840*	Requesting publication of SIR after Examiner action	
1251	110	2251	55	Extension for reply within first month	
1252	420	2252	210	Extension for reply within second month	
1253	950	2253	475	Extension for reply within third month	
1254	1,480	2254	740	Extension for reply within fourth month	
1255	2,010	2255	1,005	Extension for reply within fifth month	
1401	330	2401	165	Notice of Appeal	
1402	330	2402	165	Filing a brief in support of an appeal	
1403	290	2403	145	Request for oral hearing	
1451	1,510	1451	1,510	Petition to institute a public use proceeding	
1452	110	2452	55	Petition to revive - unavoidable	
1453	1,330	2453	665	Petition to revive - unintentional	
1501	1,330	2501	665	Utility issue fee (or reissue)	
1502	480	2502	240	Design issue fee	
1503	640	2503	320	Plant issue fee	
1460	130	1460	130	Petitions to the Commissioner	
1807	50	1807	50	Processing fee under 37 CFR 1.17(q)	
1806	180	1806	180	Submission of Information Disclosure Stmt	
8021	40	8021	40	Recording each patent assignment per property (times number of properties)	
1809	770	2809	385	Filing a submission after final rejection (37 CFR 1.129(a))	
1810	770	2810	385	For each additional invention to be examined (37 CFR 1.129(b))	
1801	770	2801	385	Request for Continued Examination (RCE)	
1802	900	1802	900	Request for expedited examination of a design application	

Other fee (specify) ______________

*Reduced by Basic Filing Fee Paid

SUBTOTAL (3) ($)

SUBMITTED BY

(Complete (*if applicable*))

Name (Print/Type)	DANIEL J. SWIRSKY	Registration No. (Attorney/Agent)	45,148	Telephone	516-620-4572 (US)
Signature				Date	MARCH 23, 2004

BASIS SELECTION IN A FINITE SET

1. INTRODUCTION

Finding a "good basis" for a vector space is a classical question in harmonic analysis and frame theory. A more restrictive question is to find a good basis from a given **finite** set of vectors, a topic treated in the signal processing area under names like "method of frames" [1], "matching pursuit" [2] and "basis pursuit" [3]. In our work we consider another variant in which the basis is chosen within a **given** finite set, and must be a "good basis" for **all** the vectors in the set.

Our work on basis selection is motivated by a sensor selection problem for interference cancellation in digital communication. This problem leads to the following question: Given a set $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ of *n* vectors in R^m, where *m<n*, find a subset *S* of size *m* which can serve as a good basis for the remaining *n-m* vectors. Hence, in our problem the goodness of the basis is evaluated relative to the **same** set from which it is selected.

To make the notion of a "good basis" concrete, consider the following model of *n* linearly distorted noisy measurements of a vector $\mathbf{u} = (u_1, u_2, \cdots, u_m)^T$:

$$x_i = \langle \mathbf{a}_i, \mathbf{u} \rangle + z_i, \quad i = 1, .., n \qquad (1)$$

where x_i is the *i*-th measurement, $\mathbf{a} = (a_{i1}, a_{i2}, \ldots, a_{im})^T$ is the corresponding vector of linear distortion coefficients, z_i is the corresponding noise, and $\langle .. \rangle$ denotes inner product.

We wish to find a subset $S \subset \{1, \ldots, n\}$ of size $|S| = m$ such that the measurements $\{x_k, k \in S\}$ are "good sensors" for the remaining measurements $\{x_i, i \notin S\}$.

We define subset goodness in two ways:

(i) Low noise amplification;

(ii) Small residual entropy.

The former notion leads to a criterion of **small expansion coefficients** of $\mathbf{a}_i$ in terms of $\{\mathbf{a}_k, k \in S\}$ i.e. the expansion coefficients vector having a small l_∞ norm, while the latter notion leads to a criterion of **maximum determinant** of the *m*x*m* matrix A_S composed of the vectors $\{\mathbf{a}_k, k \in S\}$.

We arrive at these criteria from probabilistic arguments. Assume that $u_1, u_2, \cdots, u_m$ are i.i.d. random variables ~ *N(0,1)* mutually independent of $z_1, z_2, \cdots, z_n$ which are i.i.d. ~ $N(0, \sigma^2)$. The minimum Mean Squared Error (MSE) estimate of x_i from $\{x_k, k \in S\}$, which is given in general by the conditional expectation $\overline{x_i} = E\{x_i \mid x_k, k \in S\}$, takes in this case a linear form:

$$\overline{x_i} = \langle \mathbf{g}_i, \mathbf{x}_S \rangle, \qquad (2)$$

where $g_i = (g_{i1}, g_{i2}, \cdots, g_{im})^T$ is a vector of linear estimation coefficients, and x_S is the vector with coefficients $\{x_k, k \in S\}$. Furthermore, as $\sigma^2 \to 0$ the optimal g_i approaches the Least Squares (LS) solution, i.e., the expansion of $\mathbf{a}_i$ in terms of the vectors $\{\mathbf{a}_k, k \in S\}$:

$$\mathbf{a}_i = \sum_{k \in S} g_{i,k} \cdot \mathbf{a}_k = A_S \cdot \mathbf{g}_i \ . \qquad (3)$$

Combining (1), (2) and (3), it follows that the estimation error of the LS solution is

$$\overline{x}_i - x_i = \langle \mathbf{g}_i, \mathbf{z}_S \rangle - z_i \ , \qquad (4)$$

where $\mathbf{z}_S$ is the vector $\{z_k, k \in S\}$, hence the resulting MSE is[1]

$$E[\overline{x}_i - x_i]^2 = \sigma^2 (\|\mathbf{g}_i\|^2 + 1) \ . \qquad (5)$$

We see that the noise is amplified by the expansion coefficients of $\mathbf{a}_i$ relative to the basis S. We say that a basis S is an α-amplifier if the expansion coefficients of all vectors outside S are absolutely bounded by α, i.e., for $i \notin S$

$$| g_{i,j} | \leq \alpha \qquad \text{for all } j. \qquad (6)$$

Finally, we say that a basis is good in the sense of noise amplification if it is a 1-amplifier, i.e., if $| g_{i,j} | \leq 1$ for all j.

We turn to motivate the second criterion of basis goodness. The residual entropy of the measurements relative to a basis S is defined as the conditional differential entropy of the measurements outside S given the measurements in S,

$$h(x_i, i \notin S \mid x_k, k \in S) \qquad (7)$$

(see [4] for the definition of h(..)).

This quantity determines the Shannon capacity of an n-lines vector channel, with additive noises $x_1, \ldots, x_n$, assuming lines $k \in S$ act as "sensors" (provide channel side information) for the rest of the lines. The smaller the residual entropy is, the higher is the capacity of lines $i \notin S$.

Now, by the chain rule for joint entropy we have, [4],

$$\begin{aligned} h(x_1, \ldots x_n) &= h(\mathbf{x}_k, k \in S) \\ &\quad + h(\mathbf{x}_i, i \notin S \mid \mathbf{x}_k, k \in S) \end{aligned} \qquad (8)$$

Furthermore, since $x_1, \ldots, x_n$ are jointly Gaussian,

$$e^{2h(\mathbf{x}_k, k \in S)} = | \det(A_S + \sigma^2) | \ . \qquad (9)$$

Thus, minimizing the residual entropy over the choice of S amounts to maximizing $| \det(A_S + \sigma^2) |$. As $\sigma^2 \to 0$ this becomes

$$S^* = \underset{\{S \subset \{1,\ldots,n\}, |S|=m\}}{\text{argmax}} | \det(A_S) | \ . \qquad (10)$$

As we shall see in the sequel, the two notions of goodness (unit noise amplification and maximum basis determinant) are closely related via Cramer's law. Every **locally** optimal solution for (10) (i.e., a subset such that replacing one vector does not increase its

[1] Note that this is the MSE of the LS solution for any σ^2 (not only small), and it is always an upper bound on the Baysian MMSE.

determinant) is a 1-amplifier basis. However, not every 1-amplifier basis achieves the global maximum in (10).

Geometrically, the determinant of A_S amounts to the product of the lengths of $\{\mathbf{a}_k, k \in S\}$ and the sines of the angles between each vector and the linear subspace spanned by the previous vectors (in some order). Hence, large determinant corresponds to long and close to orthogonal vectors. This partially resembles a search for the shortest basis of a given lattice.

For a lattice the basis determinant is fixed (it is the volume of the lattice basic cell), so minimizing the vectors' lengths is equivalent to making the angles as close to 90° as possible. See the LLL algorithm, [5], for an efficient search for a **reduced basis** of a lattice.

Solving (10) requires, in principle, searching all $\binom{n}{m}$ subsets and calculating their determinants. This implies $\sim n^m$ determinant calculations. On the other hand, a possible greedy solution (similar to matching pursuit [2]) sequentially selects the residual longest vector in a Gram-Schmidt-like process, implying **linear** complexity in n. However, this solution only guarantees a 2^{m-1}-amplifier basis, and a far from optimum basis determinant.

2. *Typical Results*

In this work we investigate the gap in performance between the optimum solution and low complexity variations on a greedy solution. We consider both the noise amplification and the maximum determinant basis selection criteria.

Some typical results of this work are:

1. Worst noise amplification: Applying the greedy algorithm on a given set of vectors results in a basis that is at most a 2^{m-1}-amplifier, and this bound may be achieved.
2. Efficient search for low noise amplification basis: For every set of vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$, a basis which is a $m^{1/m}$-amplifier (or better) can be found with a complexity of $O(n \cdot m^4)$.
3. Worst case determinant gap: If the absolute determinant of every concatenation of a vector of A_S with an m-1 subset of vectors is less than or equal to the absolute determinant of A_S, then the maximal determinant M is bounded by $M \leq \sqrt{m}^{m/(m+1)} \, |\det(A_S)|$, and there exist examples for which the bound is achieved.
4. For every $m>k>1$, we can construct examples of a set of vectors, where replacing any subset of k vectors from A_S by any k vectors does not increase the determinant of A_S, yet the maximal determinant M satisfies $M \geq \sqrt{m/k}^{m} \, |\det(A_S)|$.

The first two results are "optimistic" in the sense of finding low noise amplification basis with low complexity, while the latter results are "pessimistic" about the complexity of finding a basis with maximal determinant.

3. *References*

[1] I. Daubechies, *Time-frequency localization operators: a geometric phase space approach*, IEEE Trans. On Information Theory, 34(1988), pp. 605-612.

[2] S. Mallat and Z. Zhang, *Matching Pursuit in a time-frequency dictionary*, IEEE Trans. On Signal Processing, 41(1993), pp. 3397-3415.
[3] S.S. Chen, D.L. Donoho, M.A. Saunders, *Atomic decomposition by Basis Pursuit*, SIAM J. Sci Comp., 20,1 (1999) pp. 33-61.
[4] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley & Sons, New York, 1991.
[5] K. Lenstra, H. W. Lenstra, and L. Lovasz, *Factoring Polynomials with Rational Coefficients*, Math. Ann. 261, (1982) pp. 515-534.

SELF BASIS SELECTION IN A FINITE SET II

1. INTRODUCTION

Finding a "good basis" for a vector space is a classical question in harmonic analysis and frame theory [1]. A more restrictive question is to find a good basis from a given **finite** set of vectors, a topic treated in the signal processing area under names like "matching pursuit" [2] and "basis pursuit" [3].

Our work on basis selection is motivated by a sensor selection problem for interference cancellation in digital communication. This problem leads to the following question: Given a set $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ of *n* vectors in R^m, where *m*<*n*, find a subset *S* of size *m* which can serve as a good basis for the remaining *n-m* vectors. Hence, in our problem the goodness of the basis is evaluated relative to the **same** set from which it is selected
To make the notion of a "good basis" concrete, consider the following model of *n* linearly distorted noisy measurements of a vector $\mathbf{u} = (u_1, u_2, \cdots, u_m)^T$:

$$x_i = \langle \mathbf{a}_i, \mathbf{u} \rangle + z_i, \qquad i = 1, \ldots, n \tag{1}$$

where x_i is the *i*-th measurement, $\mathbf{a} = (a_{i1}, a_{i2}, \ldots, a_{im})^T$ is the corresponding vector of linear distortion coefficients, z_i is the corresponding noise, and <..> denotes inner product.
We wish to find a subset $S \subset \{1, \ldots, n\}$ of size $|S| = m$ such that the measurements $\{x_k, k \in S\}$ are "good sensors" for the remaining measurements $\{x_i, i \notin S\}$.

We define subset goodness in two ways:
(i) Low noise amplification;
(ii) Small residual entropy.
The former notion leads to a criterion of **small expansion coefficients** of $\mathbf{a}_i$ in terms of $\{\mathbf{a}_k, k \in S\}$, while the latter notion leads to a criterion of **maximum determinant** of the *mxm* matrix A_S composed of the basis vectors $\{\mathbf{a}_k, k \in S\}$.
We arrive at these criteria from probabilistic arguments. Assume that $u_1, u_2, \cdots, u_m$ are i.i.d. random variables ~ *N(0,1)* mutually independent of $z_1, z_2, \cdots, z_n$ which are i.i.d. $\sim N(0, \sigma^2)$. The minimum Mean Squared Error (MSE) estimate of x_i from $\{x_k, k \in S\}$, which is given in general by the conditional expectation $\bar{x}_i = E\{x_i \mid x_k, k \in S\}$, takes in this case a linear form:

$$\bar{x}_i = \langle \mathbf{g}_i, \mathbf{x}_S \rangle, \tag{2}$$

where $g_i = (g_{i1}, g_{i2}, \cdots, g_{im})^T$ is a vector of linear estimation coefficients, and x_S is the vector with coefficients $\{x_k, k \in S\}$. Furthermore, as $\sigma^2 \to 0$ the optimal g_i approaches the Least Squares (LS) solution, i.e., the expansion of a_i in terms of the vectors $\{a_k, k \in S\}$:

$$a_i = \sum_{k \in S} g_{i,k} \cdot a_k = A_S \cdot g_i \, . \quad (3)$$

Combining (1), (2) and (3), it follows that the estimation error of the LS solution is

$$\overline{x}_i - x_i = \langle g_i, z_S \rangle - z_i \, , \quad (4)$$

where z_S is the vector $\{z_k, k \in S\}$, hence the resulting MSE is[2]

$$E[\overline{x}_i - x_i]^2 = \sigma^2(\|g_i\|^2 + 1) \, . \quad (5)$$

We see that the noise is amplified by the expansion coefficients of a_i relative to the basis S.

We say that a basis S is an α-amplifier if the expansion coefficients of all vectors outside S are absolutely bounded by α, i.e., for $i \notin S$

$$| g_{i,j} | \le \alpha \qquad \text{for all } j. \quad (6)$$

Finally, we say that a basis is good in the sense of noise amplification if it is a 1-amplifier, i.e., if $| g_{i,j} | \le 1$ for all j.

We turn to motivate the second criterion of basis goodness. The residual entropy of the measurements relative to a basis S is defined as the conditional differential entropy of the measurements outside S given the measurements in S,

$$h(x_i, i \notin S \mid x_k, k \in S) \quad (7)$$

(see [4] for the definition of h(..)).

This quantity determines the Shannon capacity of an n-lines vector channel, with additive noises $x_1, \ldots, x_n$, assuming lines $k \in S$ act as "sensors" (provide channel side information) for the rest of the lines. The smaller the residual entropy is, the higher is the capacity of lines $i \notin S$.

Now, by the chain rule for joint entropy we have, [4],

$$h(x_1, \ldots x_n) = h(x_k, k \in S) + h(x_i, i \notin S \mid x_k, k \in S) \quad (8)$$

Furthermore, since $x_1, \ldots, x_n$ are jointly Gaussian,

---

[2] Note that this is the MSE of the LS solution for any σ^2 (not only small), and it is always an upper bound on the Baysian MMSE.

$$e^{2h(x_k, k\in S)} = |\det(A_S + \sigma^2)|. \qquad (9)$$

Thus, minimizing the residual entropy over the choice of S amounts to maximizing $|\det(A_S + \sigma^2)|$. As $\sigma^2 \to 0$ this becomes

$$S^* = \underset{\{S\subset\{1,\dots,n\}, |S|=m\}}{\operatorname{argmax}} |\det(A_S)|. \qquad (10)$$

As we shall see in the sequel, the two notions of goodness (unit noise amplification and maximum basis determinant) are closely related via Cramer's law. Every **locally** optimal solution for (10) (i.e., a subset such that replacing one vector does not increase its determinant) is a 1-amplifier basis. However, not every 1-amplifier basis achieves the global maximum in (10).

Geometrically, the determinant of A_S amounts to the product of the lengths of $\{a_k, k\in S\}$ and the sines of the angles between each vector and the linear subspace spanned by the previous vectors (in some order). Hence, large determinant corresponds to long and close to orthogonal vectors. This partially resembles a search for the shortest basis of a given lattice.
For a lattice the basis determinant is fixed (it is the volume of the lattice basic cell), so minimizing the vectors' lengths is equivalent to making the angles as close to 90° as possible. See the LLL algorithm, [5], for an efficient search for the shortest basis of a lattice.

Solving (10) requires, in principle, searching all $\binom{n}{m}$ subsets and calculating their determinants. This implies $\sim n^m$ determinant calculations. On the other hand, a possible greedy solution (as in matching pursuit [2]) sequentially selects the residual longest vector in a Gram-Schmidt-like process, implying **linear** complexity in n. However, this solution only guarantees a 2^{m-1} - amplifier basis (see Section 3), and a far from optimum basis determinant.

In this work we investigate the gap in performance between the optimum solution and low complexity variations on the greedy solution above. We consider both the noise amplification and the maximum determinant basis selection criteria. The next section describes the basis selection algorithms, while Sections 3 and 4 present our results relative to the two goodness criteria.

2. *Basis Selection Algorithms*

We shall consider two algorithms for Basis Selection.

1. Residual Longest Vector Selection.
2. One by One Replacement Algorithm.

Eventually we shall combine the 2 algorithms to get a low complexity algorithm for choosing a basis, which is close to a 1-amplifier basis.

1. Residual Longest Vector Selection.

This algorithm proceeds as follows:

Given a set $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$ of n vectors in R^m, define an m stage algorithm for choosing a set $S \subset \{1, \ldots, n\}$ of size $|S| = m$, such that $\{\mathbf{a}_k, k \in S\}$ will serve as a good basis for the given set.

Define I to be the identity mxm matrix, $Q_0 = 0$, S_0 to be the empty set.

At each stage i do the following:

a. Choose $k_i = \arg\max_{k \notin S_{i-1}} (\| (I - Q_{i-1} Q_{i-1}^T) \mathbf{a}_k \|)$. (i.e. choose the vector with maximal difference from its projection on the span of its predecessors).

b. Define S_i to be the (ordered) set of indices already chosen, i.e. $S_i = S_{i-1} \cup k_i$, and define A_{S_i} to be the mxi matrix composed of the (ordered) basis vectors $\{\mathbf{a}_k, k \in S_i\}$.

c. Compute a QR decomposition of A_{S_i}, i.e. compute $A_{S_i} = Q_i \cdot R_i$, where Q_i is a matrix of order mxi with orthonormal columns, and R_i is upper triangular of order ixi.

d. If $i<m$, return to step a., else if $i=m$, define $S = S_m, A_S = A_{S_m}$, and stop.

The principle of this algorithm is similar to that of Matching Pursuit [2], and it is a natural first choice basis. It provides some bounds on the noise amplification and on the maximal determinant, but we shall see that these bounds are quite high.

2. One by One Replacement Algorithm.

This algorithm is a simple replacement algorithm.

Given a first choice for an m subset S and the corresponding matrix A_S, the algorithm proceeds as follows:

a. Define a threshold $\alpha>1$ and a maximal number of iterations I.

b. Compute the expansion coefficients of all the vectors $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n$, relative to the given basis. (i.e. for each $\mathbf{a}_i$ compute $\mathbf{g}_i = A_S^{-1} \cdot \mathbf{a}_i$).

c. Find $G = \max_{i,j}(|g_{i,j}|)$ and the coordinates i,j for which the maximum is achieved.

d. Compare G with α. If it is greater than α then replace the j-th vector of A_S by $\mathbf{a}_i$. Update the set S and the matrix A_S. If the maximal number of iterations has not been reached, repeat the algorithm from step b.

e. If G is not greater than α, then the algorithm stops and the current basis is an α-amplifier or better.

If it is known that the maximal determinant M is bounded by $M \le D \cdot |\det(A_S)|$ for the original S, then choosing $\alpha = D^{1/(I+1)}$ guarantees that the algorithm will succeed to find a basis which is an α-amplifier or better in no more than I iterations.

Moreover we shall see that with a complexity of $O(n \cdot m^4)$ the combination of the 2 algorithms can find a basis, which is quite close to a 1-amplifier. (Note that the complexity of the full search is $O(\binom{n}{m})$ which can be considerably higher than $O(n \cdot m^4)$ for $n>>m>>1$).

3. *Noise Amplification Results*

Applying the Residual Longest Vector Selection algorithm on a given set of vectors $\mathbf{a}_1,\ldots,\mathbf{a}_n$ results in a basis, which is bounded by a 2^{m-1}-amplifier. More precisely, for each $\mathbf{a}_i$ and for $j=1,\ldots,m$, the expansion coefficients $g_{i,j}$ (equation (3)) satisfy $|g_{i,j}| \le 2^{m-j}$.

Furthermore, $|\det(A_S)| \ge M / \sqrt{m}^m$, where M denotes the maximal absolute value of all the determinants associated with a subset of m vectors of $\mathbf{a}_1,\ldots,\mathbf{a}_n$.

These results may be quite tight and there exists examples of a vector set for which the algorithm indeed results in a basis which is quite close to a 2^{m-1}-amplifier, and there exist examples for which the maximal determinant is indeed $M = \sqrt{m}^m |\det(A_S)|$.

Combining the 2 algorithms we can prove that for each I, performing at most I single replacements can reduce the basis to a basis that is a $\sqrt{m}^{m/(I+1)}$-amplifier or better. Substituting for example $I = m^2/2$ we get a $m^{1/m}$-amplifier.

The complexity of the suggested algorithm is $O(n \cdot m^4)$ and it results in a basis, which is quite close to a 1-amplifier. (Note that the complexity of the full search is $O(\binom{n}{m})$ which can be considerably higher than $O(n \cdot m^4)$ for $n>>m>>1$).

4. *Residual Entropy Results*

Residual Entropy is associated with finding a set with maximal determinant.

The results we have in this case are more "pessimistic", i.e. our results suggest that finding the maximal determinant may be a problem of high complexity.

As mentioned above, applying the Residual Longest Vector Selection algorithm on a given set of vectors $\mathbf{a}_1,\ldots,\mathbf{a}_n$ results in a basis satisfying, $|\det(A_S)| \ge M / \sqrt{m}^m$.

We can construct examples where this bound is tight. In particular for each $\varepsilon>0$, there exist examples for which the Residual Longest Vector Selection results in a 1-amplifier basis

(i.e. any replacement of only 1 vector will not increase the determinant of A_S), however the maximal determinant M satisfies $M \geq \sqrt{m/(1+\varepsilon)}^{m} \,|\det(A_S)|$.

Moreover, for every $m>k>1$, we can construct examples of a set of vectors where in addition to the above, replacing any subset of k vectors from A_S by any k vectors does not increase the determinant of A_S, yet the maximal determinant M satisfies $M \geq \sqrt{m/k}^{m} \,|\det(A_S)|$.

For $k=m$-1 we have a tighter bound. If replacing any m-1 vectors of A_S by any m-1 vectors will not increase the determinant of A_S, then the maximal determinant M is bounded by $M \leq \sqrt{m}^{m/(m+1)} \,|\det(A_S)|$, and there exist examples for which the bound is achieved.

These results suggest that finding the maximal determinant, or even a subset associated with a determinant which is close to the maximal determinant, may be a problem with high complexity, and may require an exhaustive search.

5. *References*

[6] I. Daubechies, *Time-frequency localization operators: a geometric phase space approach*, IEEE Trans. On Information Theory, 34(1988), pp. 605-612.

[7] S. Mallat and Z. Zhang, *Matching Pursuit in a time-frequency dictionary*, IEEE Trans. On Signal Processing, 41(1993), pp. 3397-3415.

[8] S.S. Chen, D.L. Donoho, M.A. Saunders, *Atomic decomposition by Basis Pursuit*, SIAM J. Sci Comp., 20,1 (1999) pp. 33-61.

[9] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley & Sons, New York, 1991.

[10] K. Lenstra, H. W. Lenstra, and L. Lovasz, *Factoring Polynomials with Rational Coefficients*, Math. Ann. 261, (1982) pp. 515-534.

Choosing a Good Base for ACTC.
Algebraic Part

INTRODUCTION:
In this paper we will introduce a concept of a good base and provide a low complexity algorithm for computing a good base from a given set of vectors.
The problem presented here and the algorithms for computing the good base are directly connected to the problem of choosing a good sensor set for ACTC.
In this paper though, we will deal only with the mathematical part.

GOOD BASE:
PseudoDefinition: Given a finite set of vectors V, a good base for *span(V)* is a subset $H \subseteq V$ such that H is a base for span(V) and every vector $v \in V$ can be expressed as a linear combination of elements of H with all the coefficients "small" (in absolute value).

Definition: Given a finite set of vectors V, and a positive scalar α, a ***α good base*** for *span(V)* is a subset $H \subseteq V$ such that H is a base for *span(V)* and every vector $v \in V$ can be expressed as a linear combination of elements of H with all the absolute values of the coefficients $\leq \alpha$.

Lemma 1: For every finite set (or even closed and bounded infinite set) $V \subseteq R^m$, there exists a α good base with $\alpha = 1$.

Proof: W.l.g. we may assume that $span(V) = R^m$. For every subset H of m vectors of V, compute the determinant, and choose a subset H with the greatest determinant. According to the famous Cramer rule when expressing any other vector $v \in V$ as a linear combination of the elements of H, the coefficients are ratios of determinants where the numerator is a determinant of a matrix generated from H by changing one of its columns and the denominator is the determinant of H.
The result follows from the maximality of *det(H).*

Lemma 1 tells us that there exists a "very good" base, but the complexity for computing that good base is very large. One has to go over all $\binom{n}{m}$ subsets of size m and compute their determinants before finding the good base.

We will now proceed to show that a naïve algorithm of choosing a "good" base for V, may result in an α good base with $\alpha \approx 2^m$.
We will suggest an alternative way for finding a α good base, with $\alpha < m$, but with bounded complexity. We will also show the dependency of α on the complexity.

The first step in the suggested algorithm is to perform a generalized QR decomposition or generalized Gram Schmidt orthogonalization process, according to the following:

1. Find the largest vector in V. Choose it as the first vector in the "good" base.
2. Normalize the vector and choose the normalized vector as the first column of Q.
3. After computing i vectors, compute the next vector as the vector whose difference from its projection on the span of the previous columns of Q is the largest. Choose that vector as the $i+1$-th vector.
4. Normalize the difference of the $i+1$-th vector from the span of the previous vectors, and choose the normalized difference vector as the $i+1$-th vector of Q.
5. Increment i by 1 and return to step 3, (unless i is greater than predefined value).

After completing this process we have a base Q, an upper triangular matrix R, with the following properties:

1. $H = QR$, where H is a matrix whose columns are the vectors chosen for the base.
2. For every $v \in V$, define y to be the vector such that $v = Qy$, then for the columns of H, we associate the columns of R, while for every $v \in V$ we see that its associated vector y has the following property :

   $\sum_{j=i}^{m} y_j^2 < r_{i,i}^2$. (where $r_{i,i}$ is the i-th diagonal element of R).

   In particular for every $v \in V$, $\quad y_i^2 < r_{i,i}^2$.

In the sequel we shall assume that the original vector set V was given in that format, i.e. the first m vectors of V are the columns of R, and in addition every $v \in V$ satisfies property 2 above.

Notations:

1. R upper diagonal matrix.
2. $r_{i,i}$, the i-th diagonal element of R.
3. A_i^v, for a matrix A and a vector v, is the matrix received by replacing the i-th column of A by the vector v.

Lemma 2. The columns of R are a α good base for the set V with $\alpha = 2^{m-1}$.

Proof: By the famous Cramer equation, the coefficients involved in expressing the vector v as a linear combination of the columns of R are ratios of determinants:

$$\left|R_i^v\right| / \left|R\right|.$$

Since $R = \prod_{i=1}^{m} r_{i,i}$, the proof will follow immediately from the following:

Lemma 3: $R_i^v \leq 2^{m-i} \prod_{i=1}^{m} r_{i,i}$.

Proof: Recall the basic definition of the determinant as a sum of products associated with permutations.

$$\det(H) = \sum_{\sigma} sign(\sigma) \prod h_{\sigma(i),i}$$

In our situation if the i-th column of R is replaced by a vector v we get the following form of matrix:

$$\begin{pmatrix} r_{1,1} & * & \cdots & * & \cdots & \cdots & \cdots & * \\ 0 & \ddots & * & * & \cdots & \cdots & \cdots & * \\ 0 & \ddots & r_{i-1,i-1} & * & \cdots & \cdots & \cdots & \vdots \\ \vdots & \ddots & 0 & * & \ddots & \cdots & \cdots & \vdots \\ \vdots & \ddots & 0 & * & r_{i+1,i+1} & \ddots & \cdots & \vdots \\ \vdots & \ddots & \vdots & * & 0 & r_{*,*} & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & 0 & \ddots & * \\ 0 & \cdots & 0 & * & 0 & 0 & 0 & r_{m,m} \end{pmatrix}$$

If we choose $\sigma(j)$ for the first i-1 columns there is only one choice that can be made, namely $\sigma(j) = j$. For all other choices of $\sigma(j)$, the result will be 0.

If we now skip the i-th column and pick $\sigma(j)$ for the columns from $i+1$ to m, we can choose $\sigma(j)$ for each column to be either $\sigma(j) = j$ or we can choose $\sigma(j)$ to equal the number from $i+1$ to j that was not yet chosen, thus for each column we have 2 valid choices of $\sigma(j)$.

Altogether, we have 2^{m-i} valid choices of σ.

The i-th column is then uniquely determined.

For each choice of σ we associate with it a product $sign(\sigma) \prod h_{\sigma(i),i}$, where each element in the product satisfies: $|h_{\sigma(i),i}| \leq |r_{i,i}|$,

thus the total product $sign(\sigma) \prod h_{\sigma(i),i}$ satisfies $\left| sign(\sigma) \prod h_{\sigma(i),i} \right| \leq |\det(R)|$.

The result follows from the fact that there are only 2^{m-i} valid permutations.

DETERMINANT BOUNDS:

In this section we will give some absolute bounds and some relative bounds on the determinants of mxm matrices whose columns are subsets of the set V satisfying all the properties discussed above. Bear in mind that we have a reference matrix, namely the matrix R above whose determinant equals: $\det(R) = \prod_{j=1}^{m} r_{j,j}$

1. Relative bounds (depending on the values of the $r_{i,i}$).

   a. For each mxm matrix M, whose columns are vectors from the subset V, we have:

$$\det(M) \le r_{1,1}^{\,m}$$

   Proof: Each column of M is a vector, whose norm is $\le r_{1,1}$, therefore the result follows.

2. Absolute Bounds (depending only on the dimension m, and the basic determinant $\det(R)$).

   a. For each mxm matrix M, which was received from R by replacing $k<m$ columns with columns from the set V, we have:

$$\det(M) \le (k+1)^{m-k}\, k!\det(R)\,.$$

   For replacing all the m columns we get:

$$\det(M) \le m!\det(R)$$

   Proof: The worst case is easily seen to be when replacing the first k columns of R. In that case the number of possible choices for $\sigma(j)$, for the last columns beginning with the $k+1$-th column is $k+1$. For the m-k columns we get $(k+1)^{m-k}$ choices. For the remaining first k columns we are left with $k!$ choices and the result follows.

   b. For each mxm matrix M, which was received from R by replacing $k<m$ columns with columns from the set V, we have:

$$\det(M) \le \sqrt{m^k m!/k!}\,\det(R)$$

   For replacing all the m columns we get:

$$\det(M) \le \sqrt{m^m}\,\det(R)$$

   Proof: Compute the determinant in rows. In each row i of the matrix, each element is no larger than $r_{i,i}$.
   The worst case happens when replacing the first k columns of R.
   In this case, the first $k+1$ rows of the new matrix can contain m non-zero elements, while from then on, the number of non-zero elements decreases by 1 for every row. Therefore the norm of each row j of the first $k+1$ rows is at most $\sqrt{m}\left|r_{j,j}\right|$.
   From the $k+2$ row down to the last the multiplicative factor decreases for each row until we get that the bound for the last row is: $\sqrt{k+1}\left|r_{m,m}\right|$.

When computing a bound on the determinant, one must always take the minimal of the above bounds, thus for a small number of replacements one must choose between the

bound $r_{1,1}{}^m$ and the bound $(k+1)^{m-k}k!\det(R)$, while for a large number of replacements the choice is between $r_{1,1}{}^m$ and $\sqrt{m^k m!/k!}\det(R)$.

Theorem: Suppose B is a bound on the determinant of a replacement matrix with $k+1$ replacements, the following algorithm will generate an α *good base* with

$$\alpha \le (B/\det(R))^{1/(k+1)}$$

while performing at most k column replacements of the matrix R.

Algorithm:

1. Denote $M_0 = R$.
2. At each stage i denote the replacement matrix of that stage as M_i. Compute all the determinants of matrices, which are replacements of *1* column of M_i and check their absolute value.
3. If the maximal absolute value of a determinant thus obtained denote $MaxM_i$ is such that $|MaxM_i/\det(M_i)| \le (B/\det(R))^{1/(k+1)}$, then choose M_i as the base, else,
4. Define M_{i+1} to be the matrix for which the maximum of the determinant was obtained and go to step 2.

Proof: It is obvious that if at a stage $i<k$, condition 3 was fulfilled then, M_i is indeed a α *good base* with $\alpha \le (B/\det(R))^{1/(k+1)}$.

If not, then at stage k we get: $|\det(M_k)| \ge \det(M_0)\cdot(B/\det(R))^{k/(k+1)}$.

If there exists a vector v whose representation in terms of the base defined by M_k includes coefficients greater than $(B/\det(R))^{1/(k+1)}$, this means that the determinant of M_k^v is greater than $\left|\det(M_k)\cdot(B/\det(R))^{1/(k+1)}\right| \ge \det(M_0)\cdot(B/\det(R)) = B$.

But B is the bound on a replacement matrix with $k+1$ replacements. Thus we cannot get an inequality in the algorithm for k successive iterations.

REMARK: In the above algorithm, one does not have to compute each determinant individually, but rather to solve the linear equation system for each of the vectors. The complexity for this process is relatively small. (linear in n (the number of vectors), cubic in *m).*

EXAMPLES:

1. For no replacements we get as large as $\alpha = 2^m$.
2. For one replacement we get $\alpha \le 3^{(m-2)/2}\sqrt{2}$.
3. For m replacements the bound gets dramatically reduced to $\alpha \le \sqrt{m}$.

4. For $m^2/2$ replacements the bound gets further reduced to $\alpha \le m^{1/m} \to 1$.

EXTENSION:
In the following section we shall generalize the problem of approximating a finite set of vectors in an m dimensional vector space by a set of $k<m$ vectors.
For that matter we need to generalize our definitions:

Definition: Given a finite set of vectors V, and positive scalars α,λ, an ***$\alpha\,\lambda$ good approximating set*** for V is a subset $H \subseteq V$ such that every vector $v \in V$ can be approximated as a linear combination of elements of H $\sum_{i=1}^{k} c_i h_i$ such that:

$$\left\{ \left\| v - \sum_{i=1}^{k} c_i h_i \right\|_2 + \lambda \max(|c_i|) \right\} < \alpha$$

The motivation for such a definition is that it leaves room for balancing between two different needs:

1. Generate a close approximation of every vector v. The first term in the inequality takes care of that.
2. Avoid noise enhancement. That is the job of the second term.

In Donoho et al there is a slightly different approach. They look at the quantity:

$$\left\{ 0.5 \left\| v - \sum_{i=1}^{k} c_i h_i \right\|_2^2 + \lambda\ sum(|c_i|) \right\} < \alpha$$

Ignoring the first term (in this case λ has no meaning), we can extend the contents of the first part of the paper to the case where the number of vectors in the dictionary we choose is smaller than the dimension of *span(V)*.

Theorem (Extension of Lemma1): For every finite set (or even closed and bounded infinite set) $V \subseteq R^m$, and every $k \le m$ there exists a subset H containing k vectors such that H is α good base with $\alpha=1$, for the set $P_H(V)$, where $P_H(V)$ is the (orthogonal) projection of V into *span(col(H))*.

Proof: Choose H such that $\det(H)$ is maximal, where $\det(H)$ for a rectangular nxm matrix with $n>m$ is defined by $\det^2(H) = \det(H^T H)$.
For every $v \in V$ solve the equation: $Hx = P(v)$.

If Q is the famous orthonormal matrix we get: $Hx = P(v) = QQ^T v$.
Multiplying by Q^T we get: $Q^T Hx = Q^T v$.
According to Cramer, each coordinate of x is given by:

$$x_i = \det((Q^T H)_i^{Q^T v}) / \det(Q^T H) = \det(Q^T H_i^v) / \det(Q^T H).$$

A generalization of Cauchy-Szhwartz would help us see that for every 2 *n*x*m* matrices *A,B*, (*n>m*),

$$\left|\det(A^T B)\right| \le \left|\det(A)\det(B)\right|.$$

For non zero determinants, equality is achieved if and only if *span(col(A)) = span(col(B))*.
Applying to our case we have: $\det(Q^T H_i^v) \le \det(H_i^v)$, and $\det(Q^T H) = \det(H)$, therefore the result follows from the maximality of det(*H*).

This proof shows that we can handle the noise enhancement problem even with a small "dictionary".

The algorithm which was defined above for *k=m*, may be used in this case also, while the determinant bounds have to be recalculated.

ESTIMATION ERROR:
Suppose $H = \{h_1, h_2, \cdots, h_{m-1}, h_m\}$ is a dictionary for the *m* dimensional set *V*.
Suppose now that we omit one vector from the dictionary, say h_m.
Consider a vector v which was originally expressed as: $v = \sum_{i=1}^{m} \alpha_i h_i$,
After the omission of h_m, only the projection on the subspace generated by the first *m-1* vectors will now be expressed as:

$$P(v) = \sum_{i=1}^{m-1} \alpha_i h_i + \alpha_m P(h_m)$$

In this case, an estimation error $\alpha_m (h_m - P(h_m))$ is generated.
The estimation error is clearly smaller when h_m is closer to its projection, or in other words when there is a sharp angle between h_m and the subspace generated by the first *m-1* vectors.

ESTIMATION ERROR vs. NOISE ENHANCEMENT:

If in the above example we express $P(h_m)$ as $P(h_m) = \sum_{i=1}^{m-1} \beta_i h_i$, then:

$$P(v) = \sum_{i=1}^{m-1} \alpha_i h_i + \alpha_m P(h_m) = \sum_{i=1}^{m-1} (\alpha_i + \alpha_m \beta_i) h_i$$

In order to ensure low noise enhancement we would like to have the coefficients β_I as small as possible. One way of keeping the coefficients low is by having a big projection error of h_m.

But this is contradictory, in some sense, to the prior requirement of a small estimation error, which will be achieved with a small projection error.

The way of finding equilibrium between the 2 is still for further study.

Choosing a Vector Set with Maximal Determinant

Consider the following questions:
Given a set $V=\{v_1, v_2, v_3, \ldots, v_n\} \subset R^m$, and assume that $span(V)= R^m$:

Choose a subset $W \subset V$, such that $W = \underset{U \subset V,\ |U|=m}{\arg\max}\left(\left|\det(U)\right|\right)$.

An exhaustive search will solve the problem for a finite V, and the complexity of such a solution will be $O(\binom{n}{m}m^3)$, since there are $\binom{n}{m}$ subsets and the complexity of each determinant computation is $O(m^3)$.

A simple low complexity solution to find a set with a determinant "close" to the maximum is a generalization of the Gram-Schmidt algorithm, which proceeds as follows:

1. Find the largest vector in V.
2. Loop:
   After choosing i vectors, the $i+1$ vector will be chosen to be the vector whose difference from its orthogonal projection on the span of the first i vectors is maximal.

The complexity of this algorithm is considerably lower than the exhaustive search.
It consists of at most m stages.
Each stage a search is done on at most n vectors.
For each search a projection computation is done which takes less than m^2 operations.
Altogether the complexity is reduced to $\sim nm^3$.

However this solution does not necessarily achieve the maximal determinant.
Some bounds on the performance of this algorithm were introduced in [1].

In this paper we analyze the following refinement of the above algorithm:

1. Choose a maximal number of iterations and denote it by *max_i*.
2. Choose a number of vectors, k, to be replaced in each iteration.
3. In each iteration up to *max_i* search over all subsets of size k of V that are not in the currently chosen set, and compute if replacing k vectors of the current set by the new vectors will increase the determinant.
4. If there is no set of size k, which can increase the determinant by replacing k vectors of the original matrix, then terminate the algorithm.

Without getting into details on the complexity of such an algorithm the following result shows that there may be a considerable difference between the maximal determinant and the determinant obtained by the algorithm:

Lemma 1: For any integers k,l with $k \leq l$, there exists an integer $m \geq l$, and a set of $2m$ vectors in R^m such that:

1. The "Gram-Schmidt" algorithm will choose the first m vectors.
2. Replacing any k vectors will not increase the determinant.
3. The maximal determinant is the determinant of the last m vectors and it is greater than the chosen determinant by a factor of $(m/k)^{m/2}$.

In particular, there exist examples of large m-s where changing $m-1$ vectors will not increase the determinant while replacing all m vectors will increase the determinant by *~sqrt(e).*

Proof:
Given $k,l,$ choose $m \geq l$ to be an integer for which a Hadamard matrix exists.

Let H_m be an m dimensional orthonormal Hadamard matrix, i.e. $H = \sqrt{1/m}\,\overline{H}$, where the elements of $\overline{H}$ are all from the set $\{\pm 1\}$.
Let $A = diag(1, 2^{-0.5}, 2^{-1}, 2^{-1.5}, \cdots, 2^{-(m-1)/2})$.
The set of $2m$ vectors, will be the columns of the matrices A and $B = \sqrt{m/k}AH$

The norm of any column of B is: $\sqrt{1/k(1+2^{-1}+2^{-2}+\cdots 2^{-(m-1)})} < \sqrt{2/k}$.
The norm of the last $m-i$ rows of every column of B is less than $\sqrt{2^{1-i}/k}$, therefore non of the columns of B would have been chosen by the Gram-Schmidt algorithm (at least for $k>2$).
Replacing any k columns of the matrix A by columns of B would not increase the determinant, since the norm of any column in a k dimensional minor of $\sqrt{m/k}H$ is at most 1, however, $\det(\sqrt{m/k}AH) = (m/k)^{m/2}\det(A)$.

For $k=m/2$, we get $\det(\sqrt{m/k}AH) = 2^{m/2}\det(A)$, so even if replacing half of the columns of A does not increase the determinant, there can still be a substantial difference if all the columns are replaced.
For $k=m-1$, this will lead to $\det(\sqrt{m/k}AH) = (m/m-1)^{m/2}\det(A) \approx \sqrt{e}\det(A)$.

CLAIMS

We claim:

1. A method substantially as described hereinabove.
2. A method substantially as illustrated in any of the drawings.
3. Apparatus substantially as described hereinabove.
4. Apparatus substantially as illustrated in any of the drawings.
5. A system substantially as described hereinabove.
6. A system substantially as illustrated in any of the drawings.